The Bombieri-Stepanov approach to the Riemann hypothesis for curves over finite fields

Raymond van Bommel

Massachusetts Institute of Technology

Simons Collaboration on Arithmetic Geometry, Number Theory, and Computation

Version of 30 November 2020

The Bombieri-Stepanov approach to the Riemann hypothesis for curves over finite fields

Raymond van Bommel

Massachusetts Institute of Technology
Simons Collaboration on Arithmetic Geometry, Number Theory, and Computation

These slides can be downloaded at
`raymondvanbommel.nl/talks/BombieriStepanov.pdf`

Introduction

Let C be a curve of genus g over a finite field $\mathbb{F}_q$. The rationality of the zeta function of C and the associated functional equation can be proven using Riemann-Roch. Indeed, the zeta function can be written as

$$Z(C, T) = \sum_{n=0}^{\infty} |\{D \in \mathrm{Div}(C) : D \text{ is effective of degree } n\}| \cdot T^n.$$

With the weak Riemann-Roch theorem rationality can be obtained, and with the strong form, including Serre duality, the functional equation can be deduced.

From this, we get that $C(\mathbb{F}_{q^e}) = q^e + 1 - \sum_{i=1}^{2g} \alpha_i^e$, where $\alpha_i \in \mathbb{C}$ are the roots of the numerator of $Z(C, T)$. In order to prove that $|\alpha_i| = \sqrt{q}$, the Riemann hypothesis for C, it suffices to prove that

$$|C(\mathbb{F}_{q^e})| = q^e + \mathcal{O}(\sqrt{q^e}).$$

This has been proved first by Weil in 1949. Weil actually gave two proofs. The first one uses intersection theory on $C \times C$ and the Hodge index theorem, and the second uses the Rosati involution on $\mathrm{Jac}(C)$.

An elementary proof by Stepanov, Smith and Bombieri

In 1969, Sergei A. Stepanov gave an elementary proof for the Riemann hypothesis for hyperelliptic curves over finite fields. This proof has been extended to the general case by Wolfgang M. Smith, and later simplified by Enrico Bombieri. This simplified proof is the topic of this talk. The main reference is:

> Enrico Bombieri, *Counting points on curves over finite fields*, 1973.

In modern language, the proof uses the polynomial method to prove that

$$|C(\mathbb{F}_{q^e})| = q^e + \mathcal{O}(\sqrt{q^e}).$$

The idea of the proof is to construct a rational function on C of small enough degree, which vanishes at almost all of the rational points in $C(\mathbb{F}_{q^e})$. This will give an upper bound on the number of points in $C(\mathbb{F}_{q^e})$. A lower bound is obtained by looking at a kind of twists of C.

Set-up

From now on, we assume that q is a square and $q > (g+1)^4$. This is not a restriction for us, as it suffices to prove the Riemann hypothesis for the base change of C to an extension of $\mathbb{F}_q$.

There is a natural bijection $\mathrm{Frob}\colon C(\overline{\mathbb{F}_q}) \to C(\overline{\mathbb{F}_q})$ induced by the Frobenius on $\mathbb{F}_q$. Let $\nu(C, \mathrm{id})$ be the number of elements of

$$\{x \in C(\overline{\mathbb{F}_q}) : \mathrm{Frob}(x) = \mathrm{id}(x)\}.$$

Of course, in this case, this is just the set $C(\mathbb{F}_q)$, so we are just counting rational points. However, this definition can be generalised by defining $\nu(C, \sigma)$ as the cardinality of

$$P(C, \sigma) := |\{x \in C(\overline{\mathbb{F}_q}) : \mathrm{Frob}(x) = \sigma(x)\}|,$$

for any automorphism σ of C. Our first goal is to prove that

$$\nu(C, \sigma) \leq q + (2g+1)\sqrt{q} + 1.$$

I will first discuss the proof for $\sigma = \mathrm{id}$, and then explain how to modify the proof for the general case.

First steps

If $P(C, \mathrm{id}) = C(\mathbb{F}_q) = \emptyset$, then there is nothing to prove, hence let $P_0 \in C(\mathbb{F}_q)$ be some element. For each positive integer m, let

$$R_m = \mathcal{L}(mP_0) = \{f \in \overline{\mathbb{F}_q}(C) : \mathrm{div}(f) + mP_0 \geq 0\}.$$

Then by Riemann-Roch we know what $m + 1 - g \leq \dim R_m \leq m + 1$, where the first inequality is an equality if $m > 2g - 2$.

On the one hand, for any integer a, let $R_\ell^{p^a} \subset R_{\ell p^a}$ be the space of functions of the shape f^{p^a} with $f \in R_\ell$. On the other hand, let $R_\ell^{\mathrm{Frob}} \subset R_{\ell q}$ be the subspace of functions of the shape $f^{\mathrm{Frob}} = f \circ \mathrm{Frob}$ with $f \in R_\ell$. Note that $f^{\mathrm{Frob}} \neq f^q$ in general, as Frob does not raise the coefficients of f to the q-th power.

Key lemma

Let a, m, and ℓ be positive integers. If $\ell p^a < q$, then the natural morphism

$$R_\ell^{p^a} \otimes_{\overline{\mathbb{F}_q}} R_m^{\mathrm{Frob}} \longrightarrow R_\ell^{p^a} R_m^{\mathrm{Frob}}$$

is an isomorphism.

Proof of key lemma.

Let $f_1, \ldots, f_r$ be a basis of R_m such that

$$\mathrm{ord}_{P_0}(f_1) < \mathrm{ord}_{P_0}(f_2) < \ldots < \mathrm{ord}_{P_0}(f_r).$$

Suppose there are $g_1, \ldots, g_r \in R_\ell$ not all equal to 0, such that

$$\sum_{i=1}^{r} g_i^{p^a} f_i^{\mathrm{Frob}} = 0. \tag{1}$$

Let i_0 is the smallest value of i such that $g_i \neq 0$. Then, on the one hand,

$$\mathrm{ord}_{P_0}(g_{i_0}^{p^a} f_{i_0}^{\mathrm{Frob}}) = \mathrm{ord}_{P_0}(g_{i_0}^{p^a}) + q \cdot \mathrm{ord}_{P_0}(f_{i_0}) \leq q \cdot \mathrm{ord}_{P_0}(f_{i_0})$$

as the function g_{i_0} is only allowed to have a pole at P_0, and hence it cannot be zero at P_0. On the other hand, for any $i > i_0$ such that $g_i \neq 0$, we have

$$\mathrm{ord}_{P_0}(g_i^{p^a} f_i^{\mathrm{Frob}}) \geq -p^a\ell + q \cdot \mathrm{ord}_{P_0}(f_i) \geq -p^a\ell + q(1 + \mathrm{ord}_{P_0}(f_{i_0})).$$

As we assumed that $\ell p^a < q$, we find that

$$\mathrm{ord}_{P_0}(g_i^{p^a} f_i^{\mathrm{Frob}}) > \mathrm{ord}_{P_0}(g_{i_0}^{p^a} f_{i_0}^{\mathrm{Frob}}), \quad \text{for all } i > i_0.$$

This contradicts equation (1). □

The next steps

Corollary

As a consequence of the key lemma, the map

$$\tau\colon \quad R_\ell^{p^a} R_m^{\mathrm{Frob}} \longrightarrow R_\ell^{p^a} R_m \colon \quad \sum g_i^{p^a} f_i^{\mathrm{Frob}} \longmapsto \sum g_i^{p^a} f_i$$

is well-defined whenever $\ell p^a < q$.

Recall that we assumed q to be a square. Now we pick a, m and ℓ such that

$$p^a = \sqrt{q}, \qquad m = \sqrt{q} + 2g, \qquad \ell = \left\lfloor \frac{g}{g+1}\sqrt{q} \right\rfloor + g + 1.$$

Then indeed

$$\ell p^a \leq \frac{g}{g+1} q + (g+1)\sqrt{q} < q,$$

as we assumed that $q > (g+1)^4$.

Claim

For this choice of a, m, and ℓ, the map τ is not injective.

Proving non-injectivity

Proof of claim.

Both m and ℓ are greater than $2g-2$. Hence, by Riemann-Roch

$$\dim(R_\ell) = \ell + 1 - g, \qquad \dim(R_m) = m + 1 - g.$$

On the other hand, $R_\ell^{p^a} R_m$ is a subspace of $R_{\ell p^a + m}$, which has dimension

$$\dim(R_{\ell p^a + m}) = \ell p^a + m + 1 - g.$$

The domain of τ has dimension $\dim(R_\ell)\dim(R_m)$, which can be shown to be greater than $\dim(R_{\ell p^a + m})$ for this choice of a, m, and ℓ. □

We find that the map τ is not injective. Let $\sum g_i^{p^a} f_i^{\mathrm{Frob}}$ be a non-zero element of the kernel of τ. Suppose that $P \in C(\mathbb{F}_q) \setminus \{P_0\}$. Then

$$f_i^{\mathrm{Frob}}(P) = f_i(\mathrm{Frob}(P)) = f_i(P).$$

Hence, $\sum g_i^{p^a} f_i^{\mathrm{Frob}}$ has a zero at P. As every element of $R_\ell^{p^a} R_m^{\mathrm{Frob}}$ is a $\sqrt{q}$-th power, the function $\sum g_i^{p^a} f_i^{\mathrm{Frob}}$ vanishes at P with multiplicity at least $\sqrt{q}$. In particular, counted with multiplicity, $\sum g_i^{p^a} f_i^{\mathrm{Frob}}$ has at least

$$\sqrt{q} \cdot (|C(\mathbb{F}_q)| - 1) \quad \text{zeros.}$$

Finishing the proof of the inequality

As a function cannot have more zeros than poles, we get the following.

Corollary

We have

$$\sqrt{q}(|C(\mathbb{F}_q)|-1) \leq \ell p^a + qm = \sqrt{q}\cdot\sqrt{q} + q\cdot(\sqrt{q}+2g),$$

or

$$|C(\mathbb{F}_q)| \leq q + (2g+1)\sqrt{q} + 1.$$

We proved that $\nu(C,1)$ satisfied the desired inequality for the upper bound, but in fact we can prove it for any $\nu(C,\sigma)$. We then need to consider a slightly different map

$$\tau : \quad R_\ell^{p^a} R_m^{\mathrm{Frob}} \longrightarrow R_\ell^{p^a} R_m^{\sigma} : \quad \sum g_i^{p^a} f_i^{\mathrm{Frob}} \longmapsto \sum g_i^{p^a} f_i^{\sigma}.$$

Then any point $P \in C(\overline{\mathbb{F}_q})$ satisfying $\mathrm{Frob}(P) = \sigma(P)$ and $P \neq P_0$ will be a zero of multiplicity at least $\sqrt{q}$ for the function in the kernel of τ. The rest of the argument is the same.

A useful proposition

Proposition

For any curve, there exists a map $C \to \mathbb{P}^1$, such that the induced field extension $\mathbb{F}_q(\mathbb{P}^1) \subset \mathbb{F}_q(C)$ is finite separable.

Proof of proposition.

Let $t \in \mathbb{F}_q(C)$ be a transcendental element, so that $\mathbb{F}_q(t) \subset \mathbb{F}_q(C)$ is finite. Now $[\mathbb{F}_q(C)^p : \mathbb{F}_q(t)^p] = [\mathbb{F}_q(C) : \mathbb{F}_q(t)]$, hence also

$$[\mathbb{F}_q(C) : \mathbb{F}_q(C)^p] = [\mathbb{F}_q(t) : \mathbb{F}_q(t)^p] = p. \tag{2}$$

Let K be the separable closure of $\mathbb{F}_q(t)$ in $\mathbb{F}_q(C)$. Then $\mathbb{F}_q(C)/K$ is purely inseparable, and as a consequence of equation (2), we have $K = \mathbb{F}_q(C)^{p^b}$ for some b. Now we can take the separable extension

$$\mathbb{F}_q(t) \subset K = \mathbb{F}_q(C)^{p^b} \overset{\text{as field}}{\cong} \mathbb{F}_q(C).$$

This extension of function fields gives a map of curves $C \to \mathbb{P}^1$ as desired. □

How to prove a lower bound?

Now our goal is to get a lower bound for $\nu(C_{\mathbb{F}_{q^e}}, \sigma)$. Let $C \to \mathbb{P}^1$ be a map of curves, such that $\mathbb{F}_q(\mathbb{P}^1) \subset \mathbb{F}_q(C)$ is separable.

First, we will assume that the extension $\mathbb{F}_q(\mathbb{P}^1) \subset \mathbb{F}_q(C)$ is Galois. In other words, we assume that $C \to \mathbb{P}^1$ is a Galois cover with Galois group G.

Lemma

As $e \to \infty$, we have

$$\sum_{\sigma \in G} \nu(C_{\mathbb{F}_{q^e}}, \sigma) = |G| \cdot |\mathbb{P}^1(\mathbb{F}_{q^e})| + \mathcal{O}(1).$$

Proof of lemma.

For any unramified point $y \in \mathbb{P}^1(\mathbb{F}_{q^e})$, there are exactly $|G|$ points above it. Each of these points is counted in exactly one of the $\nu(C_{\mathbb{F}_{q^e}}, \sigma)$. As there are only finitely many ramification points for the map $C \to \mathbb{P}^1$ and this number does not depend on e, the error caused by these ramification points goes into the $\mathcal{O}(1)$ term. □

Finishing the proof of the lower bound

The following is a corollary of the lemma and the inequalities

$$\nu(C_{\mathbb{F}_{q^e}}, \sigma) \leq q^e + \mathcal{O}(q^{e/2})$$

that we already proved.

Corollary

For each $\sigma \in G$, we have

$$\nu(C_{\mathbb{F}_{q^e}}, \sigma) = |G| \cdot (q^e + 1) - \sum_{\sigma' \neq \sigma} \nu(C_{\mathbb{F}_{q^e}}, \sigma') + \mathcal{O}(1) \geq q^e - \mathcal{O}(q^{e/2}).$$

In case $C \to \mathbb{P}^1$ is not Galois, we consider a Galois closure $D \to C \to \mathbb{P}^1$ with Galois group G. The subfield $\mathbb{F}_q(C)$ of $\mathbb{F}_q(D)$ corresponds to a subgroup H of G, and we get

$$\sum_{\sigma \in H} \nu(D_{\mathbb{F}_{q^e}}, \sigma) = |H| \cdot |C(\mathbb{F}_{q^e})| + \mathcal{O}(1),$$

analogously to the lemma. The lower bound for $|C(\mathbb{F}_{q^e})|$ now follows from the results we obtained for D.

Summary

We proved that

$$\nu(C,\sigma) \leq q + (2g+1)\sqrt{q} + 1$$

by using the polynomial method. By using the Riemann-Roch theorem and dimension counts, we constructed a rational function which vanishes with multiplicity $\sqrt{q}$ in all but one of the points in $P(C,\sigma)$. We compared the number of zeros with the maximum number of poles to obtain the inequality.

Then we used Galois theory and the upper bounds we just proved, to prove a similar lower bound for $\nu(C,\sigma)$. We first did this in the case there is a Galois cover $C \to \mathbb{P}^1$, and then we deduced the general case from this case.

Rationality and the functional equation for the zeta function follow easily from the Riemann-Roch theorem for algebraic curves. Hence, this finishes the proof of the Weil conjectures for C.

Reference: E. Bombieri, *Counting points on curves over finite fields*, 1973.